AI and Personhood

A Theological Perspective

Daniel Hackmann

What Is AI and Why Is It Called "Artificial Intelligence"?

For the purpose of simplicity and clarity, it would be good to start with a clear definition of AI. First mentioned in a summer school seminar at Dartmouth University in 1956, John McCarthy defined AI as "the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."[1] Many were not happy at that time with the notion of "artificial intelligence," as it sounded like something contrived or less than genuine, as opposed to "real intelligence." Since that early attempt at definition, the answer to the question "What is AI?" has become increasingly complex and complicated, to the point where the question is now being seen as unanswerable.[2] This has not only to do with what AI has become in its development in LLMs (Large Language Models) and generative AI, but what many proponents of AI would like it to become, namely AGI (Artificial General Intelligence) or so-called Superintelligence, far surpassing human intelligence, knowledge and reasoning. Mustafa Suleyman, CEO of Microsoft AI, stated in April of this year that AI "was a new kind of digital species" and that it is "a technology so universal, so powerful, that calling it a tool no longer captured what it could do for us."[3] Leaving aside the question as to whether Suleyman's claim is to be seen as hype or a real statement of what AI currently is, it is clear that his view is radically different than John McCarthy's original definition. The last part of Suleyman's quote raises an interesting question, namely, what is it that artificial intelligence should do, what is its purpose? While McCarthy's definition refers to tool-like qualities of AI ("using computers to understand human intelligence"), which presumably includes making machines that imitate human intelligence, Suleyman's statement points to a view of AI that far surpasses mere "toolness." What, then, does Suleyman and, by extension, Microsoft think that AI could or should do for us?

Many such optimistic and even utopian views of AI are heard in the current period between AI's "innovation trigger" and its "peak of inflated expectations," which we may soon be approaching.[4] Timnit Gebru, the founder of the Distributed AI Research Institute who formerly worked for Google and a growing number of

other voices are less enthusiastic about or outright skeptics of AI's touted abilities.[5] Referring to AI, Gebru has said, "A machine that solves all problems: if that's not magic, what is it?"[6] One reason for this skepticism has precisely to do with the question that has been there from the beginning of work with AI systems: Are machines capable of human intelligence, reasoning, and understanding? There have been some recent claims by Sam Altman, amongst others, that an AI system exhibited behavior that should be called reasonable because it seemed to have discovered a relationship between the name of a national capitol and its corresponding country on its own. But such claims have not been received with open arms, with some researchers claiming that such an AI system is just manipulating lookup tables, but not understanding the connections between what is being looked up and what is being referenced. Similar criticisms have been aimed at LLMs. LLMs are fed datasets with more text material than any human could read in a thousand lifetimes, and statistical models are very good at selecting what word is likely to be the best choice to follow another when we pose specific questions to ChatGTP or Copilot. However, the AI systems filtering and sorting the various combinations of words have no idea what those words mean. "Yes, large language models are built on math—but are they doing something intelligent with it?"[7] It may well be that the various definitions of AI have more to do with what various researchers and corporations want to see in AI and its potential to fulfill their dreams than with reality.

Beyond Artificial Intelligence

This leads to a larger issue that many have observed recently. We humans tend to recognize or infer intelligence in non-human entities, including AI systems, and humans have dreamt and written about such possibilities for millennia. The Greek myth of Talos, more than 2,000 years old, tells of a robot made by a god and tasked with protecting Crete by throwing boulders at enemy ships.[8] Stories from Hebrew folklore mention an inanimate monster, Golem, and Mary Shelley's *Frankenstein* both deal with inanimate beings that humans either create or reanimate. The human tendency to anthropomorphize things has been around for a very long time, and the dream of producing artificial humans has a much longer cultural history than the recent AI boom. Just as in the case of all past tendencies to anthropomorphize, current tendencies to do so with AI are also leading to misunderstandings regarding the nature and potential of these technologies. William Smart and Neil Richards call this tendency to anthropomorphize "the android fallacy."[9] That is, we are making a category mistake regarding the nature of these technological inventions.

Some scholars claim there is more going on here metaphysically and philosophically than just a tendency to anthropomorphize. Timnit Gebru and Émile Torres, in an article entitled "The TESCREAL Bundle: Eugenics and the Promise

of utopia through artificial general intelligence,"[10] argue that the stated goal of an overwhelming majority of high-tech firms is to build AGI,[11] not just AI, and that their goals are based on a series of ideologies which they identify as

1) Transhumanism: the central notion is that humanity can transcend itself.

2) Extopianism: libertarian version of transhumanism.

3) Singularitarianism: a type of transhumanism that emphasizes the melding of humans and machines in a coming utopia.[12]

4) Cosmism: the vision of the future that includes sentient AI, merging of machines and humans, space colonization, etc.

5) Rationalism: not to be confused with the Enlightenment version, this ideology was created in 2009 and concentrates on improving human reasoning and decision-making, mostly through AI.

6) Effective Altruism: Similar to Rationalism, but the altruists want to maximize their positive impact on the world.

7) Longtermism: combines much from the above ideologies, but places moral importance on "becoming a new posthuman species, colonizing space, controlling nature, maximizing economic productivity and creating as much value within the accessible universe as possible."[13]

Before getting into a serious discussion of these ideas, let us note here that many of these notions are ludicrous and considered by many on closer analysis to be outdated and laughable. Gedru and Torres are convinced that the ideological basis of the TESCREAL bundle is very similar to that underlying Anglo-American eugenics in the twentieth century. In addition to the overwhelming ethical issues involved with eugenics, it is also clear that the kind of future personhood advocated by the TESCREAL ideologies will only be open to a very wealthy elite.[14]

Furthermore, they argue that AGI cannot be tested and that it is being built without any clear idea of what goal it should have. It is simply assumed that AGI will be good for humankind.[15] Clearly, their reference to eugenics should raise important critical questions about where the search for AGI is heading. Rather than dealing primarily with the use of algorithms to make better and more intelligent machines, perhaps the real goal for many sponsors of this technology is indeed to upgrade or create a new type of humanity.

AGI Ideologies and Human Personhood

Although many of these ideologies sound much more like science fiction (with the emphasis on fiction) than serious positions, we must recognize that they are

attractive to many and have put people like Ray Kurzweil on the *New York Times* bestseller list. Given the extent of their influence, it is now time to consider what implications such ideas have and how they compare and contrast with other worldviews, especially those expressed in Christian theology and ethics. In this section, I will concentrate on how these discussions relate to discussions of human nature and explicitly with the Christian concept of *imago Dei*. The section following this one will look at the issues from the other way around, discussing whether AI systems could be considered to have personhood.

First of all, let us examine the bundle of TESCREAL ideologies. All of them subscribe to a kind of "improvementism," some with the emphasis on corporeal transcendence (the first four), the others with upgrading, enhancing productivity, and generally making everything better. A seemingly unlimited faith in the power of AI and a bright future pervades these worldviews, and all of them can be characterized as utopic.[16] Not only do utopias not exist, but fundamental questions can be raised about whether they could ever exist, at least based on past history. Is Ray Kurzweil's *The Singularity is Nearer*[17] founded on any real evidence? He states that computer superintelligence will already be achieved by 2029. His claims are predictions that many hold to be far-fetched at best, and are seen by many AI researchers as counterfactual. Many see "AI Hype" and "smoke and mirrors" as being a better description of the current status of AI and AGI than the predictions touted by Ray Kurzweil, Sam Altman, and Yuval Harari.[18] But whether these ideologists of transhumanism are making realistic predictions or merely marketing hype is one thing. What is more disturbing about their utopic claims about humans and the world has to do with the implications and accuracy (or, as I will claim, inaccuracy) of their description of reality.

What logical or philosophical grounds do they have for their optimism? What reason do we have to believe that humans will develop AI technologies in a benign way? Will the technologies become better and more caring if they are designed to think on their own and make complex decisions? What possible reason could we have for believing such claims? With extremely rare exceptions, every technology developed by humans has been capable of being used both for good and very evil ends. Not only are human technologies subject to good or evil usage, but they are also susceptible to good or evil design and, by inference, can embody good or evil values. This means that some technologies may be designed purposefully with evil intent, and benign usage of such technologies is, therefore, very limited or unlikely. One good example is the machine gun. Its stated and designed purpose is to rapidly kill as many humans as possible. In the case of AI systems that are enormously complex, what reason do we have to hope that benign development will be the rule and not the exception? What if these complex systems are allowed to develop themselves further? At the very least, the transhumanists who are already

celebrating the merging of AI and humans are not taking seriously the problem of evil. The problem of evil, of course, begins with those who design AI systems, that is, in the inevitably fallible nature of humans. If the major idea in transhumanism is that humans can transcend their own fallibility, it is hard to see what basis that hope could have. History has shown, again and again, precisely the opposite, that is, that human fallibility has been enhanced and multiplied in the use of destructive technologies, not transcended.

Speaking of transcendence, many of the transhumanists or singularitarians share the article of faith that the technological future will bring the elimination of death. That is, in other terms, humans will become godlike and be capable of deciding just how long they want to stay around on the earth. Describing such views, Mark O'Connell stated the following already seven years ago:

> It is their belief that we can and should eradicate aging as a cause of death; that we can and should use technology to augment our bodies and our minds; that we can and should merge with machines, remaking ourselves, finally, in the image of our own higher ideals.[19]

Yuval Harari expresses it this way:

> Humans don't die … because God decreed it, or because mortality is an essential part of some great cosmic plan. Humans always die due to some technical glitch…. [E]very technical problem has a technical solution. We don't need to wait for the Second Coming in order to overcome death.[20]

Human fallibility is only seen as a lack of technological progress, not as a fundamental limit to our self-transcendence. Given enough technological progress, we shall be like God. Harari, Kurzweil, and many transhumanists, including Julian Huxley, who came up with the position in 1957, are atheists. Some of them are wont to say that God does not exist, YET. Again, where does this overly optimistic view of humans as *homo deus* come from, and is there any evidence for it?

Interestingly, transhumanists generally deny the existence of a benevolent Creator who breathed intelligence and life into humans. A majority are materialists who believe that the guiding principle in the universe is chance, not intelligence, let alone intelligent design. Yet they have incredible confidence in the ability of humans to unlock the secret of life and upgrade themselves to creator gods.[21] There is absolutely no evidence in biological and biochemical science that humans will ever be able to create life, let alone create a god. As John C. Lennox explains, it is self-evident "that evolution did not produce life in the first place. The reason is that biological evolution, whatever it does, can only get going when life (*bios*) is already present! You cannot explain the existence of something on the basis of one of its consequences."[22] No one really knows how life initially started, and that fact is widely acknowledged today.

It should be sufficient at this point to note that the worldviews and faith exhibited by many AGI-optimists or utopians in the benevolent development of these technologies are, at best, ungrounded and, at worst, absolutely chilling.[23] If the guiding principle of human development is chance, their hope in a bright future can only be empty, if not cynical. Furthermore, let us point out that these notions are indeed metaphysically and not empirically based, and the grounds for their faith are less realistic than other options.

What implications do transhumanist notions have for the concept of personhood? One basic assumption made is that there is no given human nature or human personhood. The idea that humans are made in the *imago Dei* is rejected out of hand and replaced with the nineteenth-century notion that humans have created God, not the other way around. Our autonomous human selves are not only responsible for creating ourselves but are in a position to enhance ourselves and create and make new gods of ourselves. The "Longtermist" variety of transhumanism claims not only that this is possible but that it is our moral duty to transcend our own human nature.[24] This view of personhood claims that it is somehow intrinsic to human beings but can give no convincing account of why it should be intrinsic if its basis is material and unintelligent. Furthermore, it provides no account for the problem of evil, unless one considers evil simply to be an inferior level of development. But, as thousands of years of human development have demonstrated, technological development can certainly not be equated with virtue or leading to a more moral or ethical society.

Transhumanist notions actually go further, in that the final goal seems to be the abolition of human nature as it has been defined in the past. Human personhood in the future, enhanced by machine-generated intelligence, is considered by this worldview to be superior to mere biological existence. The Longtermist view considers it to be morally inferior to cling to traditional notions of personhood and human nature.

Now it is time to take a deeper look at the notion of *imago Dei* and contrast this notion and its associated notion of human finitude and fallibility with the prevalent ideas behind the development of AGI. *Imago Dei*, the notion that humans are created in the image of God, goes back to the beginning of the Judeo-Christian worldview. First of all, human nature or personhood is not intrinsic or evolved out of a material substrate. Human personhood is extrinsic to our material or corporeal basis. According to the book of Genesis, God gave humans specific characteristics that were Godlike, that is, similar to that of the Creator. These characteristics have typically been considered to include (not exhaustively) understanding, reasoning, consciousness, and an ability to love or care for that which is not oneself. It should also be noted that the human as *imago Dei* does not include the idea that the human is God, or a part of God. Human beings are similar, yet inferior to God and are finite, not infinite. The notion of finitude as a part of human personhood includes that notion of fallibility. Not only are humans capable of fallibility, but they also, in

fact, demonstrate failure and evil actions throughout the entire scope of history. At the very least, this initial sketch of what personhood means demonstrates not only a realistic analysis of the human condition but also points to a grounding for hope if, indeed, humans behave in accordance with the image of God. The fact that this set of notions locates the source of creativity outside of human nature itself adds humility to this notion of personhood, as opposed to the boundless pride demonstrated in the views expressed in transhumanism.

Furthermore, the notion of *imago Dei* has other advantages. If we are indeed created rather than the ultimate Creator, we should, therefore, be much more averse to the idea of "playing God." Playing God or assigning God-ness to something that is not God has, from the beginning of the Judeo-Christian tradition, been seen as idolatry, a breaking of the first and most important of the ten commandments. "Playing God," of course, means taking on a role that is diametrically opposed to the notion of *imago Dei*. Playing God can only result in injustice, evil, and destruction. The notion of *imago Dei*, therefore, has not only a positive role in assessing the importance of human personhood and life but also provides a curb or limit regarding human behavior.

One can also make the argument that attempting to erase a biological, spirit-endowed human nature and person and replace it with a "superior" transhuman version incorporating machine-like qualities is precisely that: playing God. Philosopher J. Budziszewski wrote already more than twenty years ago:

> To abolish and remake human nature is to play God. The chief objection to playing God is that someone else is God already. If He created human nature, if He intended it, if it is not the result of a blind fortuity that did not have us in mind – then we have no business exchanging it for another.[25]

Imago Dei points directly to the notion that human personhood comes directly from the intelligent and benevolent mind of God, a mind that values all human persons equally, not based on their IQ-levels or usefulness for the world's economic future. Why did the God of the universe create and endow humans with the *imago Dei*? The teleology here includes the notion of caring, of creating an I-Thou relationship, of fellowship with God.

This notion of human personhood can also be helpful in our discussion about AI and AGI. For Christians, the ultimate superhuman is already present in the God-man, Jesus. It is not necessary for us to abolish an outdated notion of humans and replace it with a god we have made. We can rest assured that the notion of personhood presented in TESCREAL ideologies is much less powerful as a description or prediction of the future than the notion of *imago Dei*. The notion of *imago Dei* includes the notion of human creativity, which is grounded in God's creativity. This creativity can be used to create new technologies that can indeed enhance and

improve human life. There are numerous examples of narrow AI that do just that. But *imago Dei* also includes an ontological limiting principle. We are limited and finite as creatures, and playing God is a violation of those principles, which are not accidental technological hurdles to be overcome but part of the design itself that encourages social interaction, humility, and care for one another, as God has done for us.

The Personhood of AI/AGI

Let us now look at personhood and AI from the other way around. To this point we have discussed various implications for human personhood in light of developments in the areas of AI systems and AGI. Many have suggested that the concept of personhood should now be applied to AI systems.[26] That is, AI systems should be seen as persons. Some of this discussion originates from various AI optimists, who argue that Artificial General Intelligence is right around the corner, and thus, we should recognize that this new intelligence is very similar, if not identical, to human personhood. This argument amounts to claiming that our personhood is an evolutionary construction and that there is no good argument other than speciesism that should prevent us from granting AI systems personhood. Others are claiming that the notion of personhood itself is flawed.[27]

This is not the place to lay out all of the current discussions regarding personhood, as a comprehensive treatment of these topics would demand an article at least as long as the present one. But for our concerns, it is important to note that many scholars are advocating the status of personhood for AI systems, and for a number of different reasons. The argument above that AI systems are either already or will be indistinguishable from human personhood in the very near future because of their rapidly developing intelligence amounts to claiming that AI personhood would simply be an extension of natural personhood. Two other arguments for AI personhood have also been voiced. One argument focuses on what AI systems now can do and whether their perceived creativity points to a type of personhood. For example,

> A computer program like a word processor does not own the text typed on it, any more than a pen owns the words that it writes. But AI systems now write news reports, compose songs, paint pictures—these activities generate value, but can and should they attract the protections of copyright law?[28]

Legal scholars tend to answer this question negatively. Still, there are already examples in China (ironically enough) where attempts have been made to protect AI systems with indirect application of copyright statutes. Some note that if AGI or even sentient AI develops, there will certainly need to be new considerations regarding the personhood of AI issues. The argument for declaring AI systems to have personhood also involves protecting the creators of AI systems. Until now, no

AI systems have been recognized as inventors, and patent law universally demands the name and address of an inventor of a patentable technology. These arguments all concern the protection of AI systems, which some believe will be enhanced by assigning them personhood, just as it tries to protect human personhood. What, of course, requires ethical and theological reflection is the assumption that human assignment of personhood to something that is nonhuman is a good strategy and what the implications of such a strategy would be. In a section of issue 55, number 2 presented in *Zygon* in June 2020 and entitled "Artificial Intelligence and Robotics: Contributions from the Science and Religion Forum," there is an article by Michael S. Burdett entitled "Personhood and Creation in an Age of Robots and AI: Can We Say 'You' to Artefacts?"[29] There he argues that a Christian theology of creation should indeed consider understanding AI "artefacts" as more than mere "things." But the authors of the forum do not go so far as to advocate personhood status across the board for AI systems.

This strategy is already being used in other areas, for instance in environmental or ecological law, where various nonhuman entities have been assigned the status of personhood. Already in 2017, after an extended legal struggle, the Whanganui River and the Te Urewera forest in New Zealand were declared to be persons.[30] The reason the status of personhood was assigned was in order to protect these natural resources in a way that their "thingness" was seen as not being able to provide. The arguments for this move were based on the concept of a legal person (as in a limited liability or GmbH company) and proposed as a new type of personhood. The hope was that the river and forest could be defended in court as persons and hence fend off environmental misuse better. By extension, such arguments could also be used to avoid the misuse of AI systems.

The other set of arguments has to do with protection *from* AI systems of the future. Some legal scholars argue that the status of personhood for AI systems would allow us legal protection because the status of personhood provides an agent that can be sued and, if necessary, punished in a way that a non-person cannot. For instance, a dog that maliciously bites a child is not held morally or ethically responsible for its behavior in a court of law. This does not mean that the dog in question would not be punished or exterminated, but the responsibility for its behavior always lies with its owner. In a scenario with a misbehaving sentient AI or AGI that had the status of personhood, it is then argued that the assigning responsibility or blame would be enhanced, as there would actually be an entity present that could be sued or punished.[31] This sounds to some like a wrongheaded strategy that could result in a complete watering down of the notion of personhood or personality. Brandeis Marshall claims that assigning AI the status of personhood is premature for various reasons, not the least of which is that we are really not clear at this point what AI or AGI is or whether AGI or superintelligence will ever exist. She recommends instead

"that we should focus first on building a social framework for AI use that protects the civil rights of all humans impacted by AI."[32]

While there may be some merit in thinking about AI personhood as analogous to legal persons, with their ability to enter contracts, be held legally responsible and able to be sued, there are also a number of distinct disadvantages that should be thought through carefully. Legal persons are able to hold property, buy and sell other legal entities (also persons!), hire and fire employees, and accumulate wealth. Are these the kinds of properties we want to assign to AI systems, and what possible reason do we have to think that these systems would use these faculties in a benevolent way? One argument against the analogy between AI systems as persons and the legal persons exemplified by companies is that the latter are made up of human persons. That is, there is a factual connection between their personhood and human personhood. In the former case, this does not obtain.

It is exactly this construction of an LLC or GmbH that provides a creepy loophole for AI systems as persons. Dr. Lance B. Eliot describes a procedure in which an AI could become an LLC or GmbH:

> Basically, a human goes ahead and forms a type of corporation commonly known as an LLC in the United States (a Limited Liability Company). The human puts in place an operating agreement that specifies the LLC will be entirely and solely governed by AI (or, if you prefer, makes reference to an "autonomous system" as an alternative phrasing). The human that founded the LLC makes sure to transfer the AI as to its originating ownership into the LLC. Finally, the human bows out of the LLC and fully dissociates themselves from the corporate entity.[33]

If this makes AI ethicists and theologians nervous, it should. Eliot states that this sleight of hand is possible not only in the USA but also with GmbHs in Germany and, Switzerland and in other countries. Various legal scholars have already affirmed that such a transformation of ownership would be possible. The result of such a magic trick is that an AI could act just as any other limited liability company does but without any human persons involved. It can hire and fire at will, accumulate wealth and property, and of course, oppress its employees, if it seems fit. The answer to the question of whether assigning personhood to AI can protect us from the excesses of Evil AI should now be quite clear. As a matter of fact, our assigning personhood to any nonhuman entity probably does nothing to that entity's inherent nature or status. The only thing that has changed is our perspective vis-à-vis that entity. Our anthropomorphizing of AI is a particularly dangerous case of what can result when we confuse our vision of what AI is and will become with reality. As Lance Eliot suggests, "Be very careful of anthropomorphizing today's AI," and he points out that Machine Learning and Deep Learning (forms of modern AI) are merely

filtering functions and "there isn't any AI today that has a semblance of common sense and nor has any of the cognitive wonderment of robust human thinking."[34] What we do know is that AI systems pick up the biases of their developers and that these developers may themselves not be aware of the biases they are passing on to the AI systems. Thus, our haste to assign personhood to AI systems is likely to fail or have negative ramifications, especially as we do not seem to know to what we might even assigning status of personhood.

In what ways might Christian ethical and theological notions be useful when considering whether to assign the status of personhood to nonhuman entities? Is something more precious if it is assigned the status of personhood even if it is not human? Or is the category of human personhood with associated human rights something that cannot or should not be assigned to other entities? Does a river in New Zealand or a lagoon in Spain have the same human rights as a human person? It is certainly the case that if I were to, for example, declare the cherry tree in my yard to have personhood, my tree's nature would not change. However, the way I perceive the tree would change. But would purposefully cutting the tree down be the equivalent of premeditated murder? Could the tree be held legally responsible for the damage it did to my house in the case of a storm? We quickly realize that notions of personhood are inescapably tied to ideas like consciousness, understanding, caring, language ability, and so forth. Anthropomorphizing nonhuman entities does not make them human by any stretch of the imagination.

In fact, the anthropomorphism of AI may compromise human dignity rather than enhance it. It is hard to imagine how feeling empathy with a tree, for example, is the equivalent of raising of the tree to human level. It seems more likely that we are lowering ourselves to the level of a tree.[35] But the ontological claim in the concept of the *imago Dei* is that God did not just pretend that human life is something like God's likeness, but that human life actually is in God's likeness. Whether we can similarly imbue nonhuman entities such as AI systems with human life as *imago homo* is symptomatic of something humans have struggled with for millennia, as we have attempted to imitate the creativity of the Creator.

Margaret Boden puts it this way:

> In a nutshell, over-reliance on computer "carers," none of which can really care, would be a betrayal of the user's human dignity.… In the early days of AI, the computer scientist Joseph Weizenbaum made himself very unpopular with his MIT colleagues by saying as much. "To substitute a computer system for a human function that involves interpersonal respect, understanding, and love," he insisted in 1976 is "simply obscene."[36]

On the other hand, the notion of *imago Dei* is all about the caring provision of a loving, Creator God. What a stark difference over against vapid transhuman ideologies.

Conclusion

In examining some of the issues surrounding AI and personhood, we have looked at the implications of our understanding of human personhood in the age of AI and whether AI systems should be considered persons. While it is clear that this whole area of research is vast and quickly developing and that there are many uncertainties even regarding the definition of AI and AGI, it is important and urgent that Christian ethicists and theologians engage not only in reflection about the direction such technologies are taking us but speak out with regard to many of the presuppositions and implications of these technologies for human personhood and life. It will simply not do to take a wait-and-see approach. Ideological, philosophical, legal, and moral arguments are being made for the necessity of accepting an essentially flawed transhuman view of the future of society and humankind. From a theological point of view, the TESCREAL ideologies are all theologies of glory. Based on unwarranted assumptions about what personhood really is, they invoke a teleology that denigrates human dignity and ends in what C.S. Lewis already in 1943 called "the abolition of man." A thoroughly Christian view of personhood with its positive insights and implications is needed now more than at any time in our lifetimes.

Daniel Hackmann *is Professor of Philosophical Theology and Academic Dean at Christ School of Theology. He received his Ph.D. in philosophical theology at the University of Iowa and his field of special interest is theology and science. In addition to his academic work, he is an entrepeneur in the area of high-tech consultancy. He lives with his wife close to Zürich, Switzerland.*

Notes

1. John McCarthy, "What is AI? / Basic Questions," http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html. Accessed July 29, 2024.
2. "Narrow AI" is perhaps easier to define as AI which enhances human capabilities, for example in the area of healthcare, where it can be used for more rapid diagnosis or guidance in neurosurgery. Here the goal does not seem to be superhuman intelligence or thinking. But even technological optimists see a need for caution here, especially in such areas as autonomous vehicles.
3. Cited in Will Douglas Heaven, "What is AI?," *MIT Technology Review*, July 10, 2024. https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/. Accessed July 29, 2024.
4. First two phases of Gartner's famous Hype Cycle research methodology.
5. Gebru was forced out of Google, mainly because of her two articles where she was critical of AI. She had co-led the Google AI Ethics team.

6. Cited in Heaven, "What is AI?," July 10, 2024. https://www.technologyreview.com/2024/07/10/1094475/what-is-artificial-intelligence-ai-definitive-guide/ Accessed July 29, 2024.

7. Ibid.

8. Adrienne Mayor, *Gods and Robots* (Princeton, NJ; Oxford, UK: Princeton University Press, 2018), p. 9.

9. N. M. Richards and W. D. Smart, "How Should the Law Think About Robots?" in *Robot Law*, eds., Ryan Calo, A. Michael Froomkin, and Ian Kerr (Cheltenham, UK; Northampton, MA: Edward Elgar Publishing, 2016), 18-21. https://www.google.com/books/edition/Robot_Law/7YpeCwAAQBAJ?hl=en&gbpv=1&dq=Robot+Law&pg=PR3&printsec=frontcover.

10. Timnit Gebru and Émile Torres, "The TESCREAL Bundle: Eugenics and the Promise of Utopia through Artificial General Intelligence," *First Monday* 29, no. 4 (April 1, 2024). https://doi.org/10.5210/fm.v29i4.13636.

11. Artificial General Intelligence. Ray Kurzweil, for example, says "The 21st century will be different. The human species, along with the computational technology it created, will be able to solve age-old problems…and will be in a position to change the nature of mortality in a post-biological future." Cited in John C. Lennox, *2084* (Grand Rapids, MI: Zondervan Reflective, 2020), 44. OpenAI defines AGI in its charter as "highly autonomous systems that outperform humans at most economically valuable work," "OpenAI charter," at https://openai.com/charter. Accessed July 15, 2024.

12. See Ray Kurzweil, *The Singularity is Near: When Humans Transcend Biology* (New York: Penguin Books, 2005) and Ray Kurzweil, *The Singularity is Nearer: When We Merge with AI* (New York: Viking, 2024).

13. Gebru and Torres, "The TESCREAL Bundle." https://firstmonday.org/ojs/index.php/fm/article/view/13636.

14. This discussion on Reddit in April 2024 reveals such concerns: https://www.reddit.com/r/singularity/comments/1cztcfa/why_would_the_elites_in_control_of_agi_do/?rdt=60199. Accessed July 16, 2024. The view expressed is that there seems to be no reason why wealthy elites should do anything benevolent with AGI. If the creators of AGI cannot be trusted to do anything benevolent with AGI, what hope do we have that AGI or a superintelligence would have a benevolent view of humanity?

15. Ibid.

16. Utopia refers by definition to something that has no place, that does not exist. It is instructive to note that boundless enthusiasm about the progress of science and society was a primary characteristic of 19th century European thought. Much of its optimism was dashed in the onset of World War I.

17. Part 2 of *The Singularity is Near*. The subtitle of the book is *When We Merge with AI*.

18. See for example, Jeffrey Funk, "Are we close to Peak AI Hype?" *Mind Matters*, July 12, 2024. https://mindmatters.ai/2024/07/are-we-close-to-peak-ai-hype/. Accessed July 29, 2024.

19. Mark O'Connell, *To Be a Machine; Adventures among Cyborgs, Utopians, Hackers and the Futurists Solving the Modest Problem of Death* (New York: Anchor, 2017), 2.

20. Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (New York: Harper, 2017), 22-23.

21. Gebru and Torres, in their conclusion, present a definition of AGI as "a system, which …seems to be an all-knowing machine akin to a 'god'" and then state, "We argue that attempting to build something akin to a god is an inherently unsafe practice." "The TESCREAL Bundle," section 8. https://firstmonday.org/ojs/index.php/fm/article/view/13636/11599. Accessed September 28, 2024.

22. John C. Lennox, *2084*, 103.

23. Gebru and Torres point out how discriminatory and racist some of the transhumanist ideas are: "The same discriminatory attitudes that animated first-wave eugenics are pervasive within the TESCREAL literature and community. For example, the Extropian listserv contains numerous examples of alarming remarks by notable figures in the TESCREAL movement. In 1996, Bostrom argued that 'Blacks are more stupid than whites,' lamenting that he couldn't say this in public without being vilified as a racist, and then mentioned the N-word (Torres, 2023a). In a subsequent 'apology' for the e-mail message, he denounced his use of the N-word but failed to retract his claim that whites are more 'intelligent' (Torres, 2023a). Also in 1996, Yudkowsky expressed concerns about superintelligence, writing: 'Superintelligent robots = Aryans, humans = Jews. The only thing preventing this is sufficiently intelligent robots.' Others worried that 'since we as transhumans are seeking to attain the next level of human evolution, we run serious risks in having our ideas and programs branded by the popular media as neo-eugenics, racist, neo-nazi, etc.' In fact, leading figures in the TESCREAL community have approvingly cited, or expressed support for, the work of Charles Murray, known for his scientific racism, and worried about 'dysgenic' pressures (the opposite of 'eugenic') (see Torres, 2023a). Bostrom himself identifies 'dysgenic' pressures as one possible existential risk in his 2002 paper, alongside nuclear war and a superintelligence takeover. He wrote: 'Currently it seems that there is a negative correlation in some places between intellectual achievement and fertility. If such selection were to operate over a long period of time, we might evolve into a less brainy but more fertile species, *homo philoprogenitus* ("lover of many offspring")' (Bostrom, 2002). More recently, Yudkowsky tweeted about IQs apparently dropping in Norway, although he added that the 'effect appears within families, so it's not due to immigration or dysgenic reproduction' — *i.e.*, less intelligent foreigners immigrating to Norway or individuals with 'lower intelligence' having more children." https://firstmonday.org/ojs/index.php/fm/article/view/13636/11599, section 4.2.

24. Longtermism "emphasizes the moral importance of becoming a new posthuman species." See Gebru and Torres, "The TESCREAL Bundle," section 4.2: https://firstmonday.org/ojs/index.php/fm/article/view/13636/11599

25. J. Budziszewski, *What We Can't Not Know: A Guide* (Dallas, TX: Spence Publishing Company, 2003), 56.

26. According to Simon Chesterman, "many of the arguments in favour of AI personality implicitly or explicitly assume that AI systems are approaching human qualities in a manner that would entitle them to comparable recognition before the law." Simon Chesterman, "Artificial Intelligence and the Limits of Legal Personality," *International and Comparative Law Quarterly* 69, no. 4 (October 2020): 831. https://www.cambridge.org/core/services/aop-cambridge-core/content/view/1859C6E12F75046309C60C150AB31A29/

S0020589320000366a.pdf/artificial-intelligence-and-the-limits-of-legal-personality.pdf Accessed July 15, 2024.

27. Jennifer Blumenthal-Barby argues that the concept of personhood is unhelpful at best and "at the worst it is harmful and pernicious." Her suggestion is that bioethicists stop using the term altogether and look for alternatives. Jennifer Blumenthal-Barby, "The End of Personhood," *The American Journal of Bioethics* 24, no. 1 (2024): 3. https://doi.org/10.1080/15265161.2022.2160515.
28. Chesterman, "Artificial Intelligence and the Limits of Legal Personality," 835.
29. Michael S. Burdett, *Zygon: Journal of Religion and Science* 55, No. 2 (June 2020): 349.
30. A good discussion of this is presented in Gwendolyn J. Gordon, "Environmental Personhood," *Columbia Journal of Environmental Law* 43, no. 1 (2018): 87ff. https://journals.library.columbia.edu/index.php/cjel/issue/view/392. Accessed July 15, 2024.
31. Interestingly enough, the notion of protection from environmental persons such as a river is not generally discussed. If the Whanganui River is a person, could it not, by the same logic as presented above also be held responsible for deaths and property damage when it floods its banks?
32. Brandeis Marshall, "No Personhood for AI" in *Patterns* 4, no. 10 (November 2023): Opening paragraph. https://www.sciencedirect.com/science/article/pii/S2666389923002453. Accessed July 29, 2024.
33. Lance Eliot, "Legal Personhood For AI Is Taking A Sneaky Path That Makes AI Law And AI Ethics Very Nervous Indeed," *Forbes*, November 21, 2022. https://www.forbes.com/sites/lanceeliot/2022/11/21/legal-personhood-for-ai-is-taking-a-sneaky-path-that-makes-ai-law-and-ai-ethics-very-nervous-indeed/. Accessed July 29, 2024.
34. Ibid.
35. For an interesting discussion of the clear distinction between us thinking about what it is like to be a bat, which can only amount to our thinking about what we would be like as a bat and what it is like for a bat to be a bat, See Thomas Nagel, "What is it Like to be a Bat?," *The Philosophical Review* 83, no. 4 (October 1974): 435-450. https://www.sas.upenn.edu/~cavitch/pdf-library/Nagel_Bat.pdf
36. Margaret Boden, "Robot Says: Whatever," *Aeon*, August 13, 2018. https://aeon.co/essays/the-robots-wont-take-over-because-they-couldnt-care-less. Accessed July 15, 2024.